



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

|  |  |  |   |
|--|--|--|---|
| (51) International Patent Classification 6 :<br><b>C12N 15/31, C07K 14/22, 16/12, C12Q 1/68, A61K 39/095, G01N 33/50</b>   |  | A2   | (11) International Publication Number: <b>WO 99/57280</b><br>(43) International Publication Date: 11 November 1999 (11.11.99) |
| (21) International Application Number: <b>PCT/US99/09346</b>   |  |  |   |
| (22) International Filing Date: 30 April 1999 (30.04.99)   |  |  |   |
| <b>(30) Priority Data:</b><br>60/083,758 1 May 1998 (01.05.98) US<br>60/094,869 31 July 1998 (31.07.98) US<br>60/098,994 2 September 1998 (02.09.98) US<br>60/099,062 2 September 1998 (02.09.98) US<br>60/103,749 9 October 1998 (09.10.98) US<br>60/103,794 9 October 1998 (09.10.98) US<br>60/103,796 9 October 1998 (09.10.98) US<br>60/121,528 25 February 1999 (25.02.99) US |  | (US). MASIGNANI, Vega [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). MORA, Marirosa [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). PETERSEN, Jeremy [US/US]; Arlington, VA (US). PIZZA, Mariagratia [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). RAPPOLI, Rino [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). RATTI, Giulio [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). SCALATO, Enzo [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). SCARSELLI, Maria [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). TETTELIN, Herve [US/US]; Gaithersburg, MD (US). VENTER, J., Craig [US/US]; Rockville, MD (US). |   |
| <b>(71) Applicants (for all designated States except US):</b> CHIRON CORPORATION [US/US]; 4560 Horton Street, Emeryville, CA 94608 (US). THE INSTITUTE FOR GENOMIC RESEARCH [US/US]; 9212 Medical Center Drive, Rockville, MD 20850 (US).  |  | <b>(74) Agent:</b> HARBIN, Alisa, A.; Chiron Corporation, Intellectual Property – R440, P.O. Box 8097, Emeryville, CA 94662-8097 (US).   |   |
| <b>(72) Inventors; and</b><br><b>(75) Inventors/Applicants (for US only):</b> FRASER, Claire [US/US]; Rockville, MD (US). GALEOTTI, Cesira [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). GRANDI, Guido [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). HICKEY, Erin [US/US]; Gaithersburg, MD  |  | <b>(81) Designated States:</b> AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, AR IPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).   |   |
| <b>Published</b><br><i>Without international search report and to be republished upon receipt of that report.</i>  |  |  |   |
| <b>(54) Title:</b> NEISSERIA MENINGITIDIS ANTIGENS AND COMPOSITIONS  |  |  |   |
| <b>(57) Abstract</b><br><p>The invention provides proteins from Neisseria meningitidis, including the amino acid sequences and the corresponding nucleotide sequences. The proteins are predicted to be useful antigens for vaccines and/or diagnostics.</p>   |  |  |   |
| <b>919 (46 kDa)</b>  |  |  |   |
| <b>A) PURIFICATION</b>   |  |  |   |
|  |  |  |   |

***FOR THE PURPOSES OF INFORMATION ONLY***

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

|    |                          |    |                                       |    |   |    |                          |
|----|--------------------------|----|---------------------------------------|----|---|----|--------------------------|
| AL | Albania                  | ES | Spain                                 | LS | Lesotho                                   | SI | Slovenia                 |
| AM | Armenia                  | FI | Finland                               | LT | Lithuania                                 | SK | Slovakia                 |
| AT | Austria                  | FR | France                                | LU | Luxembourg                                | SN | Senegal                  |
| AU | Australia                | GA | Gabon                                 | LV | Latvia                                    | SZ | Swaziland                |
| AZ | Azerbaijan               | GB | United Kingdom                        | MC | Monaco                                    | TD | Chad                     |
| BA | Bosnia and Herzegovina   | GE | Georgia                               | MD | Republic of Moldova                       | TG | Togo                     |
| BB | Barbados                 | GH | Ghana                                 | MG | Madagascar                                | TJ | Tajikistan               |
| BE | Belgium                  | GN | Guinea                                | MK | The former Yugoslav Republic of Macedonia | TM | Turkmenistan             |
| BF | Burkina Faso             | GR | Greece                                | ML | Mali                                      | TR | Turkey                   |
| BG | Bulgaria                 | HU | Hungary                               | MN | Mongolia                                  | TT | Trinidad and Tobago      |
| BJ | Benin                    | IE | Ireland                               | MR | Mauritania                                | UA | Ukraine                  |
| BR | Brazil                   | IL | Israel                                | MW | Malawi                                    | UG | Uganda                   |
| BY | Belarus                  | IS | Iceland                               | MX | Mexico                                    | US | United States of America |
| CA | Canada                   | IT | Italy                                 | NE | Niger                                     | UZ | Uzbekistan               |
| CF | Central African Republic | JP | Japan                                 | NL | Netherlands                               | VN | Viet Nam                 |
| CG | Congo                    | KE | Kenya                                 | NO | Norway                                    | YU | Yugoslavia               |
| CH | Switzerland              | KG | Kyrgyzstan                            | NZ | New Zealand                               | ZW | Zimbabwe                 |
| CI | Côte d'Ivoire            | KP | Democratic People's Republic of Korea | PL | Poland                                    |    |                          |
| CM | Cameroon                 | KR | Republic of Korea                     | PT | Portugal                                  |    |                          |
| CN | China                    | KZ | Kazakhstan                            | RO | Romania                                   |    |                          |
| CU | Cuba                     | LC | Saint Lucia                           | RU | Russian Federation                        |    |                          |
| CZ | Czech Republic           | LI | Liechtenstein                         | SD | Sudan                                     |    |                          |
| DE | Germany                  | LK | Sri Lanka                             | SE | Sweden                                    |    |                          |
| DK | Denmark                  | LR | Liberia                               | SG | Singapore                                 |    |                          |
| EE | Estonia                  |    |                                       |    |   |    |                          |

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 587>:

a147.seq

```

1 ATGCGACGAG AAGCCAAAAT GGCACAAACT ACACCTCAAAC CCATTGTTT
51 ATCAATTCTT TTAATCAACA CACCCCTCCT CTCCAAGCG CATGGAACGT
101 AGCAATCAGT GGGCTTGGAA ACGGTCAAGCG TCGTCGGCAA AAGCCGTCCG
151 CGGCCCACTT CGGGCTGCT GCACACTTCT ACCGCCTCCG ACAAAATCAT
201 CAGCGGCGAC ACCTTGCAC AAAAAGCCGT CAACCTGGGT GATGCTTAG
251 ACGGCGTACC GGGCATTCAAT GCCTCGCAAT ACGGGGGCGG CGCATCCGCT
301 CCCGTTATTC GCGGTCAAAC AGGCAGACGG ATTAAAGTGT TGAACCATCA
351 CGGGGAAACG GGCACATGG CGGACTTCTC TCCAGACCAT GCAATCATGG
401 TGGACAGCGC CTTGTCGAA CAGGTGCAAA TCCTGCGCGG TCCGGTTACG
451 CTCTTGACAG GCTCGGGCAA TGTGGCGGGG CTGGTCGATG TTGCGGATGG
501 CAAAATCCCC GAAAAAAATGC CTGAAAACGG CGTATCGGGC GAACTCGGAT
551 TGCCTTGAG CAGCGGCAAT CTGGAAAAAC TCACGTCGGG CGGCATCAAT
601 ATCGGTTTGG GCAAAACACT TGTATTGCAAC ACGGAAGGGC TGTAACCGCAA
651 ATCGGGGAT TACGCGTAC CGCGTTACCG CAATCTGAAA CGCCTGCCCG
701 ACAGCCACGC CGATTCGAA ACGGGCAGCA TCGGGCTGTC TTGGGTTGGC
751 GAAAAAGGCT TTATCGGCG AGCATACAGC GACCGTCGCG ACCAATATGG
801 TCTGCCTGCC CACAGCCACG AATACGATGA TTGCCACGCC GACATCATCT
851 GGCAAAAGAG TTTGATTAAC AAACGCTATT TGCAGCTTTA TCCGCACCTG
901 TTGACCGAAG AAGACATCGA TTACGACAAT CCGGGCTTGA GCTGCGGCTT
951 TCACGACGAC GATGATGCAC ACGCCCATGC CCACAAACGGC AAACCTTGGA
1001 TAGACCTGCG CAACAAACGC TACGAACCTC GCGCCGAATG GAAGCAACCG
1051 TTCCCCGGTT TTGAAGCCCT GCGCGTACAC CTGAACCGCA ACGACTACCG
1101 CCACGACGAA AAAGCAGGGCG ATGCACTGAG AAACCTTTTT ACAACACCAA
1151 CGCAAAACGC CCGTATCGAG TTGCGCCACC AACCCATAGG CCGTCTGAAA
1201 GGCAGCTGGG GCGTGCATAA TTGGGACAA AAATCCAGTG CTTTATCTGC
1251 CACATCCGAA CGCGTCAAAAC AACCGATGCT GCTTGACAAT AAAGTGCAC
1301 ATTACAGCTT TTTCGGTGTA GAACAGGCAA ACTGGGACAA CTTCACGCTT
1351 GAAGGGGGCG TACGCGTGGAA AAAACAAAAA GCCTCCATCC GCTACGACAA
1401 AGCATTGATT GATCGGGAAA ACTACTACAA CCATCCCCTG CCCGACCTCG
1451 GCGCGCACCG CCAAACCGCC CGCTCATTG CACTTTCGGG CAACTGGTAT
1501 TTCACGCCAC AACACAAACT CAGCCTGACC GCCTCCCATC AGGAACGCC
1551 GCCGTCAACG CAAGAGCTGT ACGCACACGG CAAACACGTC GCCACCAACA
1601 CCTTTGAAGT CGGCAACAAA CACCTCAACA AAGAGCGTTC CAACAATATC
1651 GAACTCGCGC TGGGCTACGA AGGCAGCCGC TGGCAATACA ATCTGGCACT
1701 CTACCGAAC CGCTTCGGCA ACTACATTCA CGCCCAAAAC TTAAACGACG
1751 GACCGGGCCC CAAATCCATC GAAGACGACA CGGAAATGAA GCTCGTGGC
1801 TACAACCAAT CCGGTGCGGA CTTCTACGGC GCGGAAGGCG AAATCTACTT
1851 CAAACCGACA CCGCGCTTACCG GCATCGGCGT TTCCGGCGAC TATGTACGAG
1901 GCCGTCTGAA AAACCTGCCT TCCCTACCCG GCAGGGAAAGA CGCCTACGGC
1951 AACCGCCAC TCATTGCCCA AGCCGACCAA AACGCCCTC GCGTTCCGGC
2001 TGCGCGCTC GGCCTCCACC TGAAAGCTC GCTGACCGAC CGCATCGATG
2051 CCAATTGGA CTACTACCGC GTGTTCGCCC AAAACAAACT CGCCCGCTAC
2101 GAAACCGACA CGCCCGGACA CCATATGCTC AACCTCGCG CAAACTACCG
2151 CCGCAATACG CGCTATGGCG AGTGGAAATG GTACGTCAAAC GCGACAAACC
2201 TGCTCAACCA ATCCGTTTAC GCCCACAGCA GCTTCCTCTC TGATACGCGC
2251 CAAATGGGCC CGAGCTTACCG CGGCGCGTG AACGTGAAGT TTAA

```

This corresponds to the amino acid sequence <SEQ ID 588; ORF 147.a>:

a147.pep

```

1 MRREAKMAQT TLKPIVLSIL LINTPLLSQA HGTEQSVGLE TVSVVGKSRP
51 RATSGLLHTS TASDKIISGD TLRQKAVNLG DALDGVPGIH ASQYGGGASA
101 PVIKGQTGRR IKVLNHHGET GDMADFSRPH AIMVDSALSQ QVEILRGPV
151 LLYSSGNVAG LVDVADGKIP EKMPENGVSG ELGLRLSSGN LEKLTSGGIN
201 IGLGKNFVLH TEGLYRKSGD YAVPRYRNKL RLPDSHADSDQ TGSIGLSWVG
251 EKGFIGAAYS DRRDQYGLPA HSHEYDDCHA DIIWOKSLIN KRYLQLYPHL
301 LTEEDIDYDN PGLSCGFHDD DDAHAHAHNG KPWIDLRLNKR YELRAEWKQP
351 FPGFEALRVH LNRNDYRHDE KAGDAVENFF NNQTQNARIE LRHQPIGRILK
401 GSWGVQYLGQ KSSALSATSE AVKQPMLLDN KVQHYSFFGV EQANWDNFTL
451 EGGVRVEKQK ASIRYDKALI DRENYYNHPL PDLGKAHRQTA RSFALSGNWY
501 FTPQHKLSLT ASHQERLPST QELYAHGKHV ATNTFEVGNK HLNKERSNNI
551 ELALGYEGDR WQYNLALYRN RFGNYIYAQT LNDGRGPKSI EDDSEMKLVR

```

1101 AGGTCGTCTG AAAGGCAGCT GGGCGTGCA ATATTTACAA CAAAAATCCA  
 1151 GTGCTTATC TGCCATATCC GAAGCGTTA AACAAACCGAT GCTGCTTGAC  
 1201 AACAAAGTGC AACATTACAG CTTTTTCGGT GTAGAACAGG CAAACTGGGA  
 1251 CAACTTCACG CTTGAAGGAG GCGTACCGT GGAAAACAA AAAGCCTCCA  
 1301 TTCAGTACGA CAAAGCATTG ATTGATCGGG AAAACTACTA CAACCACCCC  
 1351 CTGCCCGACC TCGGCACGCA CCGCCAAACC GCCCCTCAT TCGCACTTTC  
 1401 GGGCAACTGG TATTTCACGC CACAACACAA ACTCAGCCTG ACCGCCTCCC  
 1451 ATCAGGAACG CCTGCCGTCA ACGCAAGAGC TGTACGCACA CGGCAAACAC  
 1501 GTGCCACCA ACACCTTGA AGTCGGCAC AAACACCTCA ACAAAAGAGCG  
 1551 TTCCAACAAAT ATCGAACCTG CGCTGGCTA CGAAGGGCAG CGCTGGCAAT  
 1601 ACAAATCTGC ACTCTACCGC AACCGCTTCG GTAACTACAT TTACGCCAA  
 1651 ACCTTAAACG ACGGACGCGG CCCCAAATCC ATCGAAGACG ACAGCGAAAT  
 1701 GAAGCTCGTG CGCTACAAACC AATCGGCAG CGACTTCTAC GGCGCGGAAG  
 1751 GCGAAATCTA CTTCAAACCG AACCGCGCT ACCGCATCGG CGTTTCCGGC  
 1801 GACTATGTAC GAGGCCGTCT GAAAAACCTG CCTTCCTAC CGGGCAGAGA  
 1851 AGATGCCTAC GGCAACCGTC CTTTCATCGC ACAGGACGAC CAAAATGCC  
 1901 CCCGTGTTCC GGCTGCGCAG CTCGGCTTCC ACCTGAAAGC CTCGCTGACC  
 1951 GACCGTATCG ATGCCAATTT GGACTACTAC CGCGTGTTCG CCCAAACAA  
 2001 ACTCGCCCGC TACGAAACGC GCACGCCCGG ACACCATATG CTCAACCTCG  
 2051 GCGCAAACTA CCGCCGCAAT ACGCGCTATG GCGAGTGGAA TTGGTACGTC  
 2101 AAAGCCGACA ACCTGCTCAA CCAATCCGTT TACGCCACCA GCAGCTTTCT  
 2151 CTCTGATAACG CCGCAAATGG GCCGCAGCTT TACCGCGGC GTGAACGTGA  
 2201 AGTTTTAA

This corresponds to the amino acid sequence <SEQ ID 586; ORF 147>:

**m147.pep** (partial)

|     |   |
|-----|---|
| 1   | ..PHKTEQSVDL ETVSVVGKSR PRATSGLLHT STASDKIISG DTLRQKAVNL  |
| 51  | GDALDGVPGI HASQYGGGAS APVIRGQTGR RIKVLNHHGE TGDMADFS      |
| 101 | HAIMVDTALS QQVEILRGPV TLLYSSGNVA GLVDVADGKI PEKMPENGVS    |
| 151 | GELGLRLSSG NLEKLTSIGGI NIGLGKNFVL HTEGLYRKSG DYAVPRYRN    |
| 201 | KRLPDSHADS QTGSIGLSWV GEKGFIGVAY SDRRDQYGLP AHSHEYDDCH    |
| 251 | ADIIWQKSLI NKRYLQLYPH LLTEEDIDYD NPGLSCGFHD DDNAHAHTHS    |
| 301 | GRPWIIDLRLNK RYELRAEWKQ PFPFGFEALRV HLNRNDYRHD EKAGDAVENF |
| 351 | FNNQTQNARI ELRHQPIGRL KGSWGVQYILO QKSSALSAIS EAVKQPM      |
| 401 | NKVQHYSFFG VEQANWDNFT LEGGVRVKEQ KASIQYDKAL IDRENYYNHP    |
| 451 | LPDLGAHRQI ARSFALSGNW YFTPQHKLSL TASHQERLPS TQELYAHGKH    |
| 501 | VATNTFEVGN KHLNKERSNN IELALGYEGD RWQYNLALYR NRFGNYIYAQ    |
| 551 | TLNDGRGPKS IEDDSEMKLV RYNQSGADFY GAEGEIYFKP TPRYRIGVSG    |
| 601 | DYVRGRKLKNL PSLPGREDAY GNRPFIAQDD QNAPRVPAAAR LGFHLKASLT  |
| 651 | DRIDANLDYY RVFAQNKLAR YETRTPGHM LNMGANYRRN TRYGEWNWYV     |
| 701 | KADNLLNQSV YAHSSFLSDT PQMGRSFTGG VNVKF*                   |

Computer analysis of this amino acid sequence gave the following results:  
Homology with a predicted ORF from *N. gonorrhoeae*

**m147 / g147** 92.3% identity in 142 aa overlap

|                 |   |
|-----------------|---|
| <b>m147.pep</b> | 10            20            30  |
|                 | PHKTEQSVDL ETVSVVGKSR PRATSGLLHTS<br>  :      :      :      :                                       |
| <b>g147</b>     | MRREAKMAQITLKPIVL SILLINTPLLAQAHETEQSVGLET VSVVGKSR PRATSGLLHTS                                     |
|                 | 10            20            30            40            50            60                            |
| <b>m147.pep</b> | 40            50            60            70            80            90                            |
|                 | TASDKIISGDTLRQKAVNLGDALDGVPGIHASQYGGGASAPVIRGQTGRRIKVLNHHGET<br>      :      :      :      :      : |
| <b>g147</b>     | TASDKIISGDTLRQKAVNLGDALDGVPGIHASQYGGGASAPVIRGQTGRRIKVLNHHGET  |
|                 | 70            80            90            100            110            120                         |
| <b>m147.pep</b> | 100            110            120            130            140            150                      |
|                 | GDMADFS PDHAIMVDTALSQQVEILRGPV TLLYSSGNVAGLVDVADGKIPEKMPENGVSG<br>      :      :      :      :      |
| <b>g147</b>     | GDMADFS PDHAIMVDTALSQQVEILRGPV TLLYSSGNVAGAGQCCRWNPKNA  |

601 YNQSGADFYG AEGEIYFKPT PRYRIGVSGD YVRGRLKNLP SLPGREDAYG  
 651 NRPLIAQADQ NAPRVPAARL GVHLKASLTD RIDANLDYYR VFAQNKLARY  
 701 ETRTPGHHML NLGANRYRRNT RYGEWNWYVK ADNLINQSVY AHSSFLSDTP  
 751 QMGRSFTGGV NVKE\*

m147/a147 98.1% identity in 734 aa overlap

|          |  | 10           | 20  | 30         |        |
|----------|--|--------------|-----|------------|--------|
| m147.pep |  | PHKTEQSVDLET | VS  | VVGKSRPRAT | SGLHTS |
| a147     | MRREAKMAQTTLKPIVLSILLINTPLLSQAHGTEQSVGLET  |              |     |            |        |
|          |  | 10           | 20  | 30         | 40     |
|          |  | 50           |     | 60         |        |
| m147.pep | TASDKIISGDTLQRQKAVNLGDALDGVPGIHASQYGGGASAP | 40           | 50  | 60         | 70     |
| a147     | VIRGQTGRRIKVLNHHGET                        |              |     |            |        |
|          |  | 80           |     | 90         |        |
| m147.pep | GDMADFS PDHAIMVDTALSQQVEILRGPVTL           | 100          | 110 | 120        | 130    |
| a147     | LYSSGNVAGLVDVADGKIP                        |              |     |            |        |
|          |  | 140          |     | 150        |        |
| m147.pep | PEKMPENGVSG                                |              |     |            |        |
| a147     | TASDKIISGDTLQRQKAVNLGDALDGVPGIHASQYGGGASAP | 130          | 140 | 150        | 160    |
|          | VIRGQTGRRIKVLNHHGET                        |              |     |            |        |
|          |  | 170          |     | 180        |        |
| m147.pep | ELGLRLSSGNLEKLTS                           | 160          | 170 | 180        | 190    |
| a147     | GGINIGLGKNFVLHTEGLYRKSGDYAVPRYRNLKRLPDSHAD |              |     |            |        |
|          | SQ   | 200          |     | 210        |        |
| m147.pep | ELGLRLSSGNLEKLTS                           | 190          | 200 | 210        | 220    |
| a147     | GGINIGLGKNFVLHTEGLYRKSGDYAVPRYRNLKRLPDSHAD |              |     |            |        |
|          | SQ   | 230          |     | 240        |        |
| m147.pep | TGSIGLSWVGEKGFIGVAYS                       | 220          | 230 | 240        | 250    |
| a147     | DERRDQYGLPAHSHEYDDCHADI                    |              |     |            |        |
|          | IWQKS                                      | 260          |     | 270        |        |
| m147.pep | LINKRYLQLYPHL                              | 250          | 260 | 270        | 280    |
| a147     | TGSIGLSWVGEKGFIGAAYS                       | 290          |     | 300        |        |
|          | DRRDQYGLPAHSHEYDDCHADI                     |              |     |            |        |
|          | IWQKS                                      | 290          |     | 300        |        |
| m147.pep | LTEEDIDYDNPGLSCGF                          | 280          | 290 | 300        | 310    |
| a147     | HDDDNAHAHTHSGRPWIDL                        |              |     |            |        |
|          | RNKRYELRAEWKQPFPGFEALRVH                   | 320          |     | 330        |        |
| m147.pep | LTEEDIDYDNPGLSCGF                          | 310          | 320 | 330        | 340    |
| a147     | HDDDAHAHANGKPWIDL                          |              |     |            |        |
|          | RNKRYELRAEWKQPFPGFEALRVH                   | 350          |     | 360        |        |
| m147.pep | LNRNDYRHDEKAGDAVENF                        | 340          | 350 | 360        | 370    |
| a147     | FNNQTNARIELRHQPI                           |              |     |            |        |
|          | GLKGSGWVQYLQKSSALS                         | 380          |     | 390        |        |
| m147.pep | 370  | 380          | 390 | 400        | 410    |
| a147     | 370  |              | 380 |            | 390    |
| m147.pep | AVKQPMLLDNKVQHYSFF                         | 400          | 410 | 420        | 430    |
| a147     | GVQEANWDNFTLEGGVR                          |              |     |            |        |
|          | VEKOKASI                                   | 440          |     | 450        |        |
| m147.pep | 430  | 440          | 450 | 460        | 470    |
| a147     | 430  |              | 440 |            | 450    |
| m147.pep | PDLGAHRQTARS                               | 460          | 470 | 480        | 490    |
| a147     | FALSGNWYFTPQHKL                            |              |     |            |        |
|          | SLTASHQERLPSTQELYAHGKH                     | 500          |     | 510        |        |
| m147.pep | 500  | 510          | 520 | 530        | 540    |
| a147     | 500  |              | 510 |            | 520    |
| m147.pep | HLNKERSNNIELALGYEG                         | 520          | 530 | 540        | 550    |
| a147     | DRWQYNLALYRNRF                             |              |     |            |        |
|          | FGNYIYAQTLNDGRGP                           | 560          |     | 570        |        |
| m147.pep | 550  |              | 560 |            | 570    |
| a147     | 550  |              | 560 |            | 570    |

|          | 550  | 560 | 570 | 580 | 590 | 600 |
|----------|--|-----|-----|-----|-----|-----|
| m147.pep | 580  | 590 | 600 | 610 | 620 | 630 |
|          | YNQSGADFYGAEGEIYFKPTPRYRIGVSGDYVRGRLKNLPLPGREDAYGNRPFIAQDDQ    |     |     |     |     |     |
| a147     | 610  | 620 | 630 | 640 | 650 | 660 |
|          | YNQSGADFYGAEGEIYFKPTPRYRIGVSGDYVRGRLKNLPLPGREDAYGNRPLIAQADQ    |     |     |     |     |     |
| m147.pep | 640  | 650 | 660 | 670 | 680 | 690 |
|          | NAPRVPAA RLGFHLKASLTDRIDANLDYYRVFAQNKLARYETRTPGHHMLNLGAN YRRNT |     |     |     |     |     |
| a147     | 670  | 680 | 690 | 700 | 710 | 720 |
|          | NAPRVPAA RLGVHLKASLTDRIDANLDYYRVFAQNKLARYETRTPGHHMLNLGAN YRRNT |     |     |     |     |     |
| m147.pep | 700  | 710 | 720 | 730 |     |     |
|          | RYGEWNWYVKADNLLNQS VYAHSSFLSDTPQMGRSFTGGVNVKFX                 |     |     |     |     |     |
| a147     | 730  | 740 | 750 | 760 |     |     |
|          | RYGEWNWYVKADNLLNQS VYAHSSFLSDTPQMGRSFTGGVNVKFX                 |     |     |     |     |     |

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 589>:

g148.seq

|     |            |             |            |            |            |
|-----|------------|-------------|------------|------------|------------|
| 1   | ATGGCGTTAA | AAACATCAAA  | CTTGGAACAC | GCAATGctgg | ttcaTCCCGA |
| 51  | AgctATgagt | gtccgcgcgc  | TTGccgAcaa | AATCCGCAAA | AtcgaAAact |
| 101 | gGCCGCAAAA | AGgcataCTTA | TTCCACGACA | TCACGCCCGT | CCTGCAAAGT |
| 151 | GCGGAATACT | TCCGCCTTTT  | GGTCGATTG  | CTGGTTTACC | GCTATATGGA |
| 201 | TCAGAAAATC | GACATCGTTG  | CCGGCTTGG  | CGCCGCGGGC | TTCATTATCG |
| 251 | GCGCGGACT  | CGCCTACCA   | CTCAaCGtcg | gctTCGTCCC | CATCCGCAAA |
| 301 | AAAGGCAAGC | TGCCTTTGA   | AACCGTATCG | CAAAGCTAcg | cgcTCGAATA |
| 351 | CGGGGAAGCT | GCGGTGGAAA  | TCCACACCGA | tgccgTCAAA | CCCGGTTCGC |
| 401 | GCGTCTTGCT | GGTCGATGAT  | TTGGTTGCCA | CGGGCGGCAC | AATGCTTGCC |
| 451 | GGGCTGGAAC | TGATCCGCAA  | ACTCGGGGG  | GAAATTGTCG | AAggcgccgC |
| 501 | CATTTGGAA  | TTTACCGACC  | TTCAAGGGGG | CAAGAATATC | CGCGCAAGTG |
| 551 | GCGCGCCCTT | ATTTACCCCTG | CTTCAAAACG | AAGGCTGCAT | GAAAGGCTGA |

This corresponds to the amino acid sequence <SEQ ID 590; ORF 148.ng>:

g148.pep

|     |            |            |                   |                   |            |
|-----|------------|------------|-------------------|-------------------|------------|
| 1   | MALKTSNLEH | AMLVHPEAMS | VGALADKIRK        | IENWPQKGIL        | FHDITPVLQS |
| 51  | AEYFRLLVDL | LVYRYMDOKI | <u>DIVAGLDARG</u> | <u>FIIGAALAYQ</u> | LNVGFVPIRK |
| 101 | KGKLPFETVS | QSYALEYGEA | AVEIHTDAVK        | PGSRVLLVDD        | LVATGGTMLA |
| 151 | GLELIRKLGG | EIVEAAAILE | FTDLOGGKNI        | RASGAPLFTL        | LQNEGCMKG* |

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 591>:

m148.seq

|     |            |             |            |            |            |
|-----|------------|-------------|------------|------------|------------|
| 1   | ATGGCGTTAA | AAACATCAAA  | CTTGGAACAC | GCAATGCTGG | TTCATCCCGA |
| 51  | AGCTATGAGT | GTCCGCGCGC  | TTGCCGACAA | AATCCGCAAA | ATCGAAAAC  |
| 101 | GGCCGCAAAA | AGGCATCTTA  | TTCCACGACA | TCACGCCCGT | CCTTCAAAGC |
| 151 | GCGGAATACT | TCCGCCTTTT  | GGTTGATTTA | TTGGTTTACC | GCTATATGGA |
| 201 | TCAGAAAATC | GACATCGTTG  | CCGGTTTGG  | CGCCGCGGGC | TTCATTATCG |
| 251 | GCGCGGACT  | CGCCTACCA   | CTCACGTCG  | GTTTCGTCCC | CATCCGCAAA |
| 301 | AAAGGCAAGC | TGCCTTTGA   | AACCGTATCG | CAAAGCTACG | CGCTCGAATA |
| 351 | CGGGGAAGCT | GCGGTGGAAA  | TCCACACCGA | TGCCGTCAA  | CTCGGTTCGC |
| 401 | GCGTCTTGCT | GGTCGATGAT  | TTGATTGCCA | CGGGCGGCAC | GATGCTTGCC |
| 451 | GGACTGGAAC | TGATCCGCAA  | ACTCGGGGG  | GAAATTGTCG | AAggcgccgC |
| 501 | CATTTGGAA  | TTTACCGACC  | TTCAAGGGGG | CAAGAATATC | CGTGCAAGCG |
| 551 | GCGCGCCCTT | ATTTACCCCTG | CTTCAAAACG | AAGGCTGTAT | GAAAGGCTGA |

This corresponds to the amino acid sequence <SEQ ID 592; ORF 148>:

m148.pep

|    |            |            |                   |                   |            |
|----|------------|------------|-------------------|-------------------|------------|
| 1  | MALKTSNLEH | AMLVHPEAMS | VGALADKIRK        | IENWPQKGIL        | FHDITPVLQS |
| 51 | AEYFRLLVDL | LVYRYMDOKI | <u>DIVAGLDARG</u> | <u>FIIGAALAYQ</u> | LNVGFVPIRK |

351 LRVHLNRNDY HHDEKAGDAV ENFFNNKTHN ARIELRHQPI GRLKGSGWVQ  
 401 YLGQKSSALS AIPETVQQPM LIIDNNVRHYS FFGVEQANWD NFTLEGGVRY  
 451 EKQKASIRYD KALIDRENYY NQPLPDGLAH RQTARSFALS GNWYFTPQHK  
 501 LSLTASHQER LPSTQELYAH GKHVATNTFE VGNKHLNKER SNNIELALGY  
 551 EGDRWQYNLA AYRNRFQNYI YAQLNDGRG PKSIEDDSEM KLVRYNQSGA  
 601 DFYGAEGEYY FKPTPRYRIG VSGDYVRGRRL KNLPSLPGRE DPYGKRPFIA  
 651 QADQNAPRIP AARLGFLKLT SLTDRIDANL DYRYRVFAQNK LARYETRTPQH  
 701 HHMLNLGAN YRRNTRYGEWN WYVKADNLLN QSVYAHSSFL SDTPQMGRSF  
 751 TGGVNVKF\*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 603>:

m149-1.seq

```

1 ATGGCACAAA CTACACTCAA ACCCATTGTT TTATCAATT CTTTAATCAA
51 CACACCCCTC CTCGCCAAG CGCATGAAAC TGAGCAATCG GTGGATTG
101 AAACGGTCAG CGTCGTCGGC AAAAGCCGTC CGCGGCCAC GTCGGGGCTG
151 TTGCACACTT CGACCCCTC CGACAAAATC ATCTCGGGG ATACCTTGCG
201 CCAAAAAGCC GTCAACTTGG GCGACGCTT AGACGGCGTA CGGGGCATCC
251 ACGCTTCGCA ATACGGCGC GGCACGTCTG CTCCCGTCAT TCGCGGTCAA
301 ACAGGCAGGC GGATTAAGT GTGAAACCAT CACGGCGAAA CAGGCAGATAT
351 GGCAGATTTC TCGCCCGATC ACGCCATTAT GGTAGATACC GCCTTGTGCG
401 AACAGGTGCA AATCTGCGC GGGCCGGTTA CGCTCTTGTG CAGCTCGGGC
451 AATGTGGCGG GGCTGGTCGA TTGTCGGAT GGCAAAATCC CCGAAAAAAAT
501 GCCTGAAAAC GGCATCGG CGCAACTCGG ATTGCGTTG AGCAGCGGCA
551 ATCTGGAAA ACTCACGTCC GGCAGCATCA ATATCGGTTT GGGCAAAAAC
601 TTTGTATTGC ACACGGAAAG CCGTGTACCGC AAATCGGGGG ATTACGCCGT
651 ACGCGTGTAC CGCAATCTGA AACGCCCTGCG CGACAGCCAC GCCGATTTCG
701 AAACGGGCAG CATCGGGCTG TCTTGGGTTG GCGAAAAAGG TTTTATCGGC
751 GTAGCGTACA GCGACCGTCG CGACCAATAT GGTCTGCCG CCCACAGCCA
801 CGAACATCG GATTGCCACG CGCACATCAT CTGGCAAAAG AGCTTGATTA
851 ACAAACGCTA TTTACAGCTT TATCCGCACC TGTTGACCGA AGAACACATC
901 GATTGACACA ATCCGGGCTT GAGCTGCGGC TTCCACGACG ACGATAATGC
951 ACACGCACAC ACCCACAGCG GCAGACCGTG GATAGACCTG CGCAACAAAC
1001 GCTACGAAC CCCTGCCGAA TGAAGCAAC CGTCCCCCGG TTTTGAAGCC
1051 CTGCGCGTAC ACCTGAACCG CAACGACTAC CGCCACGACG AAAAAGCAGG
1101 CGATGCGTC GAAAACCTTT TTAACAAACCA AACGCAAAAC GCCCGCATCG
1151 ATTTGCGCCA CAAACCCATA GGTGCTCTGA AAGGCAGCTG GGGCGTGC
1201 TATTTAACAC AAAATCCAG TGCTTTATCT GCCATATCG AAGCGTTAA
1251 ACAACCGATG CTGCTGACA ACAAAAGTGCA ACATTACAGC TTTTCGGTG
1301 TAGAACAGGC AAAACTGGGAC AACTTCACGC TTGAAGGAGG CGTACCGTG
1351 GAAAACACAA AAGCCCTCCAT TCAGTACGAC AAAGCATTGA TTGATCGGG
1401 AAACACTACT AACACCCCCC TGCCCCGACCT CGGCGCGCAC CGCCAAACCG
1451 CCCGCTCATC CGCACCTTCG GGCACACTGGT ATTTACGCG ACAACACAAA
1501 CTCAGCCTGA CGCCTCCCA TCAGGAACGC CTGCGCTCAA CGCAAGAGCT
1551 GTACGCACAC GGCAAAACACG TCGCCACCAA CACCTTGAA GTCGGCAACA
1601 AACACCTCAA CAAAGAGCGT TCCAACAAATA TCAGACTCGC GCTGGGCTAC
1651 GAAGGCAGC GCTGGCAATA CAATCTGGCA CTCTACCGCA ACCGCTTCGG
1701 TAACTACATT TACGCCAAA CCTTAAACGA CGGACGCGGC CCCAAATCCA
1751 TCGAAGACGA CAGCGAAATG AAGCTCGTGC GCTACAACCA ATCCGGCGCC
1801 GACTTCTACG GCGCGGAAGG CGAAATCTAC TTCAAACCGA CACCGGGCTA
1851 CCGCATCGC GTTTCGGCG ACTATCTACG AGGCGCTCTG AAAAACCTGC
1901 CTTCCTTACCC CGCGCAGAGAA GATGCCCTACG GCAACCGTCC TTTCATCGA
1951 CAGGACGACC AAAATGCCCG CCGTGTCTCG GCTGCGCGCC TCGCTTCCA
2001 CCTGAAAGCC TCGCTGACCG ACCGTATCGA TGCAATTG GACTACTACC
2051 GCGTGTTCGC CAAAAACAAA CTCGCCCCTG ACAGAACCGG CACGCCCGGA
2101 CACCATATGC TCAACCTCG CGCAAACACTAC CGCCGCAATA CGCGCTATGG
2151 CGAGTGGAAAT TGGTACGTCA AAGCCGACAA CCTGCTAAC CAATCCGTTT
2201 ACGCCACAC CAGCTTCTC TCTGATACGC CGCAAATGGG CGCGACCTTT
2251 ACCGGCGGCG TGAACGTGAA GTTTAA
```

This corresponds to the amino acid sequence <SEQ ID 604; ORF 149-1>:

m149-1.pep

```

1 MAQTTLKP ILSILLINTPL LAQAHETEQS VDLETVSVVG KSRPRATSGL
51 LHTSTASDKI ISGDTLRQKA VNLDALDGV PGIHASQYGG GASAPVIRGQ
101 TGRRIKVLNH HGETGDMADF SPDHAIMVDT ALSQQVEILR GPVTLLYSSG
151 NVAGLVLVDAD GKIEPKMPEN GVSGELGLRL SSGNLEKLTS GGINIGLGN
201 FVLHTEGLYR KSGDYAVPRY RNKLRLPDSH ADSQTGSIGL SWVGEKGFIG
251 VAYSDRRDQY GLPAHSHEYD DCHADIWQK SLINKRQLQ YPHLLTEEDI
301 DYDNPGGLSCG FHDDDNAH AH THSGRPWIDL RNKRYELRAE WKQPFPGFEA
351 LRVHLNRNDY RHDEKAGDAV ENFFNNQTNQ ARIELRHQPI GRLKGSGWVQ
401 YLQQKSSALS AISEAVKQPM LLDNKVQHYS FFGVEQANWD NFTLEGGVRY
451 EKQKASIRYD KALIDRENYY NHPLPDGLAH RQTARSFALS GNWYFTPQHK
501 LSLTASHQER LPSTQELYAH GKHVATNTFE VGNKHLNKER SNNIELALGY
```

|          |     |   |     |     |     |     |
|----------|-----|---|-----|-----|-----|-----|
| a149     |     | PAARLGVLKASLTDRIDANLDYRVFAQNKLARYETRTPGHMLNLGANRNRNTRYGEW |     |     |     |     |
|          | 250 | 260   | 270 | 280 | 290 | 300 |
| m149.pep | 310 | 320   | 330 | 340 |     |     |
| a149     |     | NWYVKADNLLNQSVDYAHSSFLSDTPQMGRSFTGGVNVKFX                 |     |     |     |     |
|          | 310 | 320   | 330 | 340 |     |     |

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 601>:

**g149-1.seq**

```

1 ATGGCACAAA TCACACTCAA ACCCATTGTT TTATCAATT CTTTAATCAA
51 CACACCCCTC CTCGCCAAG CGCATGAAAC TGACCAATCG GTGGGCTTGG
101 AAACGGTCAG CGTCGTCGGC AAAAGCCGTC CGCGCGGCAC TTGGGGCTG
151 CTGCACACTT CGACCGCCTC CGACAAAATC ATCTCCGGCG ATACTTTGG
201 CCAAAAAAGCC GTCAACTTGG GCGACGCTTT GGACGGCGTA CGGGCATCC
251 ACGCTTCGCA ATACGGCGGC GGCGCATCCG CTCCCGTTAT TCGGGTCAA
301 ACGGGCAGAC GGATTAAGT ATTGAACCAT CACGGCGAAA CGGGCGATAT
351 GGCGGACTTT TCTCCCGATC ACGCCATTAT GGTAGATACC GCCTTGTGCG
401 AACAGGGTGA AATCCTGCGC GGGCCGGTTA CGCTCTTGTA CAGCTCGGC
451 AATGTGGCGG GGCTGGTCGA TGTTGCCGAT GGAAAATCC CCGAAAAAAT
501 GCCTGAAAAC GGCGTATCGG GCGAagccgg ATTGCCTTG AGCAGCGGCA
551 ATTTAGAAAA ACTGACATCC CGAGGCATCA ATATCGGACT GGGCAAAAAC
601 TTCGTGCTG ATACCGAAGG CTTGTACCGC AAATCGGGCG ATTACGCCGT
651 ACCGCCTTAC CGCAATCTGA AACGCCCTGCC CGACAGCCAT GCCGATTGCG
701 AAACGGGCAG CATCGGGCTG TCTTGGGTGG GCGAAAAAAG CTTTATCGGC
751 GCAGCATACA GCGACCGTCG CGACCGCTAC GGCGTCCCTG CCCACAGCCA
801 CGAATACGAT GATTGCCACG CCGACATCAT CTGGCAAAAG AGTTTGATCA
851 ACAAAACGCTA TTGCACTT TATCCGACT TGTGACCGA AGAAGACATC
901 GATTACGACA ATCCGGCTT GAGCTGCGGC TTCCACGAGC GCGACGGTGC
951 ACACGCACAC ACCCACAACG GCAAAACCGTG GATAGACTG CGCAACAAAC
1001 GCTACGAACCT CGCGCCGAA TGGAAAGCAGC CATTCCCCGG TTTTGAAGCC
1051 CTGCGCGTAC ATCTGAACCG CAATGACTAC CACCAACGAGC AAAAACGAGG
1101 CGATGCACTA GAAAACCTCT TCAACAAACAA AACACACAAC GCCCCTATCG
1151 AGTTGCGCCA CCAACCCATA GGCGCTCTGA AAGGCAGCTG GGGCGTGCAA
1201 TATTTGGGAC AAAAATCCAG CGCGCTTTCC GCGATTCCCG AAACCGTCCA
1251 ACAACCGATG TTGATTGACA ACAATGTCCG CCATTACAGC TTTTCCGGT
1301 TAGAACAGGC AAATTGGGAC AACTTCACGC TTGAAGGGGG CGTACCGGTG
1351 GAAAAACAAA AAGCCTCCAT CGGGTACGAC AAAGCATTGA TTGATCGAGA
1401 AAACACTAC AACACGCCCT TGCCCGGACCT CGCGCGGCAC CGCCAAACCG
1451 CCCGCTCGTT CGCACCTTCG GGCAACTGTT ATTCACGCC ACACCAACAA
1501 CTCAGCCTGA CGCCTCCCA TCAGGAACGC CTGCGTCAA CGCAAGAACT
1551 GTACGACACG GCGAACGACG TCGCCACCAA CACCTTTGAA GTCGGCAACA
1601 AACACCTCAA CAAAGAGCGT TCCAACAAATA TCGAACTCGC GCTGGGCTAC
1651 GAAGGGCACC GCTGGCAATA CAATCTGGCA GCCTACCGCA ACCGATTGCG
1701 CAACTACATT TACGCCAAA CCTTAAACGA CGGACGGGC CCCAAATCCA
1751 TCGAAGACGA CAGCGAAATG AAGCTCGTGC GCTACAAACCA ATCCGGTGC
1801 GACTTCTACG CGCGGGAAGG CGAAATCTAC TTCAAACCGA CACCGCGCTA
1851 CGCATCGGT GTTTCGGCG ACTATGTACG AGGCGCTCTG AAAAACCTGC
1901 CGTCCCTACC CGCGAGGGAA GATCCCTACG GCAAAACGTCC CTTCATCGCA
1951 CAAGCCGACC AAAACGCCCT CGCGATTCCG GCTGCGGCC TC GGCTTCCA
2001 CCTGAAAACC TCGCTAACCG ACCGTATCGA TGCCAATTG GACTACTACC
2051 GCGTGTTCGC CCAAAACAAA CTCGCCGCT ACGAAACCGG TAGGCCCGGA
2101 CACCATATGC TCAACCTCGG TGCAAACATAC CGCCGCAATA CGCGCTATGG
2151 CGAGTGGAAAT TGGTACGTCA AAGCCGACAA CCTGCTAAC CAATCGTTT
2201 ACGCCCCACAG CAGCTTCCCTC TCTGTACGCC CGCAAATGGG CGCAGCTTT
2251 ACCGGCGCG TAAACGTGAA GTTTTAA

```

This corresponds to the amino acid sequence <SEQ ID 602; ORF 149-1.ng>:

**g149-1.pep**

```

1 MAQITLKPIV LSILLINTPL LAQAHETEQS VGLETVSVVG KSRPRATSGL
51 LHTSTASDKI ISGDTLRQKA VNLGDALDGV PGIHASQYGG GASAPVIRGO
101 TGRRIKVLNH HGETGDMADF SPDHAIMVDT ALSQQVEILR GPVTLLYSSG
151 NVAGLVWDAD GKIKEKMPEN GVSGEAGRLI SSGNLEKLTS AGINIGLGN
201 FVLHTEGLYR KSGDYAVPRY RNLKRLPDSH ADSQTGSIGL SWVGEKGFIG
251 AAYSDRDRY GLPAHSHEYD DCHADIIWQK SLINKRYLQL YPHLLTEEDI
301 DYDNPGISCG FHDGDGAHAH THNGKPWIDL RNKRYELRAE WKQPFPGFEA

```

551 EGDRWQYNLA LYRNRFQNYI YAQTLDNGRG PKSIEDDSEM KLVRYNQSGA  
601 DFYGAEGEYI FKPTPLRYRIG VSGDYVRGRL KNLPSLPGRE DAYGNRPFIA  
651 QDDQNAPRVP AARLGFLHLKA SLTDRIDANL DYRRVFAQNK LARYETRTPG  
701 HHMLNLGQY RRNTRYGEWN WYVKADNLLN QSVYAHSSFL SDTPQMGRSF  
751 TGGVNVKF\*

m149-1/g149-1 96.2% identity in 758 aa overlap

|            |  |     |     |     |     |     |     |
|------------|--|-----|-----|-----|-----|-----|-----|
| m149-1.pep | MAQTLKPIVLSILLINTPLAQAHETEQSVVDLETSVVVGKSRPRATSGLLHTSTASDKI      | 10  | 20  | 30  | 40  | 50  | 60  |
| g149-1     | MAQTLKPIVLSILLINTPLAQAHETEQSVVGLETSVVVGKSRPRATSGLLHTSTASDKI      | 10  | 20  | 30  | 40  | 50  | 60  |
| m149-1.pep | ISGDTLRQKAVNLGDALDGVPGIHASQYGGGASAPVIRGQTGRRIKVLNHHGETGDMADF     | 70  | 80  | 90  | 100 | 110 | 120 |
| g149-1     | ISGDTLRQKAVNLGDALDGVPGIHASQYGGGASAPVIRGQTGRRIKVLNHHGETGDMADF     | 70  | 80  | 90  | 100 | 110 | 120 |
| m149-1.pep | SPDHAIMVDTALSQQVEILRGPVTLLYSSGNVAGLVDVADGKIKEPKMPENGVSGEGLRL     | 130 | 140 | 150 | 160 | 170 | 180 |
| g149-1     | SPDHAIMVDTALSQQVEILRGPVTLLYSSGNVAGLVDVADGKIKEPKMPENGVSGEAGLRL    | 130 | 140 | 150 | 160 | 170 | 180 |
| m149-1.pep | SSGNLEKLTSAGINIGLKNFVLUHTEGLYRKSGDYAVPRYRNLKRLPDSHADSQTGSIGL     | 190 | 200 | 210 | 220 | 230 | 240 |
| g149-1     | SSGNLEKLTSAGINIGLKNFVLUHTEGLYRKSGDYAVPRYRNLKRLPDSHADSQTGSIGL     | 190 | 200 | 210 | 220 | 230 | 240 |
| m149-1.pep | SWVGEKGFIGVAYSDRRDQYGLPAHSHEYDDCHADIWIWQKSLINKRQLQYPHLLTEEDI     | 250 | 260 | 270 | 280 | 290 | 300 |
| g149-1     | SWVGEKGFIGAAYSDRRDYGLPAHSHEYDDCHADIWIWQKSLINKRQLQYPHLLTEEDI      | 250 | 260 | 270 | 280 | 290 | 300 |
| m149-1.pep | DYDNPGLSCGFHDDDNAHAHTHSGRPWIDLRLRNKRYELRAEWKQPFPFGFEALRVHILNRNDY | 310 | 320 | 330 | 340 | 350 | 360 |
| g149-1     | DYDNPGLSCGFHDDDNAHAHTHSGRPWIDLRLRNKRYELRAEWKQPFPFGFEALRVHILNRNDY | 310 | 320 | 330 | 340 | 350 | 360 |
| m149-1.pep | RHDEKAGDAVENFFNNQTQNARIELRHQPIGRLKGSWGVQYLQQKSSALSAISEAVKQPM     | 370 | 380 | 390 | 400 | 410 | 420 |
| g149-1     | RHDEKAGDAVENFFNNKTNAARIELRHQPIGRLKGSWGVQYLQQKSSALSAIPETVQOPM     | 370 | 380 | 390 | 400 | 410 | 420 |
| m149-1.pep | LLDNKVQHYSFFGVEQANWDNFTLEGGVRVEKQKASIYDALKALIDRENNYNHPLPDLGAH    | 430 | 440 | 450 | 460 | 470 | 480 |
| g149-1     | LIDNNVRHYSFFGVEQANWDNFTLEGGVRVEKQKASIYDALKALIDRENNYNQPLPDLGAH    | 430 | 440 | 450 | 460 | 470 | 480 |
| m149-1.pep | RQTARSFALSGNWYFTPQHKLSLTASHQERLPSTQELYAHGKHVATNTFEVGNKHLNKER     | 490 | 500 | 510 | 520 | 530 | 540 |
| g149-1     | RQTARSFALSGNWYFTPQHKLSLTASHQERLPSTQELYAHGKHVATNTFEVGNKHLNKER     | 490 | 500 | 510 | 520 | 530 | 540 |
| m149-1.pep | SNNIELALGYEGDRWQYNLALYRNRFGNYYIAQTLNDGRGPKSIEDDSEMKLVRYNQSGA     | 550 | 560 | 570 | 580 | 590 | 600 |
| g149-1     | SNNIELALGYEGDRWQYNLALYRNRFGNYYIAQTLNDGRGPKSIEDDSEMKLVRYNQSGA     | 550 | 560 | 570 | 580 | 590 | 600 |
| m149-1.pep | DFYGAEGEIFYFKPTPRYRIGVSGDYVRGRLKNLPSLPGREDAYGNRPFIAQDDQNAPRVP    | 610 | 620 | 630 | 640 | 650 | 660 |
| g149-1     | DFYGAEGEIFYFKPTPRYRIGVSGDYVRGRLKNLPSLPGREDAYGNRPFIAQADQNAPRIP    | 610 | 620 | 630 | 640 | 650 | 660 |

|            |  |     |     |     |     |     |
|------------|--|-----|-----|-----|-----|-----|
|            | 670  | 680 | 690 | 700 | 710 | 720 |
| m149-1.pep | AARLGFLKASLTDRIDANLDYYRVFAQNKLARYETRTPGHMLNLGANYRRNTRYGEWN<br>       :       :       :       :       :       :       :       : |     |     |     |     |     |
| g149-1     | AARLGFLKTSLTDRIDANLDYYRVFAQNKLARYETRTPGHMLNLGANYRRNTRYGEWN<br>670        680        690        700        710        720       |     |     |     |     |     |
|            | 730  | 740 | 750 | 759 |     |     |
| m149-1.pep | WYVKADNLLNQSVDYAHSSFLSDTPQMGRSFTGGVNVKFX<br>       :       :       :       :       :       :       :                           |     |     |     |     |     |
| g149-1     | WYVKADNLLNQSVDYAHSSFLSDTPQMGRSFTGGVNVKFX<br>730        740        750  |     |     |     |     |     |

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 605>:

a149-1.seq

```

1 ATGGCACAAA CTACACTCAA ACCCATTGTT TTATCAATT CTTTAATCAA
51 CACACCCCTC CTCTCCCAAG CGCATGGAAC TGAGCAATCA GTGGGCTTGG
101 AAACGGTCAG CGTCGTCGGC AAAAGCCGT CGCGCGCCAC TTCGGGGCTG
151 CTGCACACTT CTACCGCCCTC CGACAAAATC ATCAGCGGCG ACACCTTGCG
201 ACAAAAAGGCC GTCAACTTGG GTGATGCTTT AGACGGCGTA CGGGGCATTG
251 ATGCCCTCGCA ATACGGCGGC GGCGCATCCG CTCCCCTTAT TC CGGGTCAA
301 ACAGGCAGAC GGATTAAGT GTGAACCAT CACGGCGAAA CGGGCGACAT
351 GGCAGACTTC TCTCCAGACC ATGCAATCAT GGTGGACAGC GCCTGTGCG
401 AACAGGTCGA AATCCTGCGC GGTCCGGTTA CGCTCTTGTA CAGCTCGGGC
451 AATGTGGCGG GGCTGGTCGA TGTTGGCGAT GGCAAAATCC CGGAAAAAAAT
501 GCGCTGAAAC GGCCTATCGG GCGAACCTGG ATTGCGTTTG AGCAGCGGCA
551 ATCTGGAAAA ACTCACGTC GGCGGCATCA ATATCGGTTT GGGCAAAAC
601 TTTGTATTGC ACACGGAAGG GCTGTACCGC AAATCGGGGG ATTACGCGT
651 ACCCGCTTAC CGCAATCTGA AACCGCTGCG CGACAGCCAC GCCGATTGCG
701 AAACGGGCAG CATCGGGCTG TCTTGGGTTG GCGAAAAAAGG CTTTATCGGC
751 GCAGCATACA GCGAACCGTCG CGACAAATAT GGTCTGCGT CCCACAGCCA
801 CGAACATGAT GATTGCCACG CCGACATCAT CTGGCAAAAG AGTTGATTA
851 ACAAAAGCTA TTTGCAGCTT TATCCGCACC TGTTGACCGA AGAAAGACATC
901 GATTAGACA ATCCGGGCTT GAGCTGCGGC TTTCACGACG ACGATGATGC
951 ACACGCCCCAT GCCACAACG CGAACACCTTG GATAGACCTG CGCAACAAAC
1001 GCTACGAACT CGCGCCGAA TGGAAGCAAC CGTTCCCCCG TTTTGAAGCC
1051 CTGCGCTGAC ACCTGAAACG CAACGACTAC CGCCACGAGC AAAAAGCAGG
1101 CGATGCAGTA GAAAACTTTT TTAACAAACCA AACGCAAAAC GCCCCTATCG
1151 AGTTGCGCCA CCAACCCATA GGCCTGCTGA AAGGCAGCTG GGGCGTCAA
1201 TATTGGGGAC AAAAATCCG TGCTTTATCT GCCACATCGG AAGCGGTCAA
1251 ACAACCGATG CTGCTTGACA ATAAAGTGCAC ACATTACAGC TTTTTCGGTG
1301 TAGAACAGGC AAAACTGGGAC AACTTCACGC TTGAAGCGGG CGTACGCGTG
1351 GAAAACAAA AAGCCTCCAT CCGCTACGAC AAAGCATGAA TTGATCGGG
1401 AAAACTACTAC ACCATCCCC TGCCCGACCT CGGCGCGCAC CGCCAAACCG
1451 CCCGCTCATT CGCAACTTTCG GGCAACTGGT ATTTCACGCC ACAACACAAA
1501 CTCAGCTGA CCCGCCCTCCCA TCAGGAACCC CTGCCGTCAA CGCAAGAGCT
1551 GTACGCAACAC GGCAAAACAGC TGCCACCAA CACCTTGAA GTCGGCAACA
1601 AACACCTCAA CAAAGAGCGT TCCAACAATA TCGAACCTCGC GCTGGGCTAC
1651 GAAGGGCACC GCTGGCAATA CAATCTGGCA CTCTACCGCA ACCGCTTCGG
1701 CAACTACATT TACGCCAAA CCTTAAACGCA CGGACGGGGC CCCAAATCCA
1751 TCGAAGACGA CAGCGAAATC AGACTCGTC GCTACAACCA ATCCGGTGGC
1801 GACTTCTACG GCGCGGAAGG CGAAATCTAC TTCAACCCGA CACCGCGCTA
1851 CCGCATCGGC GTTCCGGCG ACTATGTACG AGGCCGCTG AAAAACCTGC
1901 CTTCCCTACC CGGCAGGGAA GACGCCCTACG GCAACCGCCC ACTCATTGCC
1951 CAAGGCCGACC AAAACGCCCG TCGCGTCCCG GCTGCCGCC TC CGGGCGTCCA
2001 CCTGAAAGCC TCGCTGACCG ACCGCATCGA TGCCAATTG GACTACTACC
2051 GCGTGTTCGC CCAAAACAAA CTCGCCCCGCT ACGAAACCGC CACGCCCGGA
2101 CACCATATGC TCAACCTCGG CGCAAACCTAC CGCCCAATA CGCGCTATGG
2151 CGAGTGGAAAT TGGTACGTCA AAGCCGACAA CCTGCTAAC CAATCCGTTT
2201 ACGCCCCACAG CAGCTTCCCTC TCTGATACGC CGCAAATGGG CGCGAGCTTT
2251 ACCGGCGGCG TGAACGTGAA GTTTTAA

```

This corresponds to the amino acid sequence <SEQ ID 606; ORF 149-1.a>:

a149-1.pep

```

1 MAQTTLKPIV LSILLINTPL LSOAHGTEQS VGLETVSVVG KSRPRATSGL
51 LHTSTASDKI ISGDTLRQKA VNLDALDGV PGIHASQYGG GASAPVIRGQ
101 TGRRRIKVNLH HGETGDMADF SPDHAIVMDS ALSQQVEILR GPVTILYSSG
151 NVAGLVDVAD GKIPKEKMPEN GVSGELGLRL SSGNLEKLTG GGINIGLGKN
201 FVLHTEGLYR KSGDYAVPRY RNLKRLPDSH ADSQTGSIGL SWVGEKGFIG
251 AAYSDRDRDQY GLPAHSHEYD DCHADIWQK SLINKRYLQL YPHLILTEEDI
301 DYDNPGLSCG FHDDDDAHAH AHNGKPWIDL RNKRYELRAE WKQPFPGFEA

```

351 LRVHLNRNDY RHDEKAGDAV ENFFNNQTQN ARIELRHQPI GRLKGSGWGVQ  
 401 YLGQKSSALS ATSEAVKQPM LLDNKVQHYS FFGVEQANWD NFTLEGGVRV  
 451 EKQKASIRYD KALIDRENYY NHPLPDGLAH RQTARSFALS GNWYFTPQHK  
 501 LSLTASHQER LPSTQELYAH GKHVATNTFE VGNKHLNKER SNNIELALGY  
 551 EGDRWQYNLA LYRNRFGNYI YAQTLNDGRG PKSIEDDSEM KLVRYNQSGA  
 601 DFYGAEGEIIY FKPTPRYRIG VSGDYVRGRL KNLPSLPGRE DAYGNRPLIA  
 651 QADQNAPRVP AARLGVHLKA SLTDRIDANL DYYRVAQNK LARYETRTPG  
 701 HHMLNLGANY RRNTRYGEWN WYVKADNLLN QSVYAHSSFL SDTPQMGRSF  
 751 TGGVNVKF\*

a149-1/m149-1 98.0% identity in 758 aa overlap

|            |  |   |  |   |     |     |     |
|------------|--|---|--|---|-----|-----|-----|
|            | 10   | 20  | 30   | 40  | 50  | 60  |     |
| a149-1.pep | MAQTTLKPIVLSILLINTPLLSQAHQTEQSVDLETSVVVGKSRPRATSGLLHTSTASDKI   | :     :     :     :     :     :     :     :     :     :     : | MAQTTLKPIVLSILLINTPLLAQAHQTEQSVDLETSVVVGKSRPRATSGLLHTSTASDKI   | :     :     :     :     :     :     :     :     :     :     : |     |     |     |
| m149-1     |  | 10  | 20   | 30  | 40  | 50  | 60  |
|            |  |   |  |   |     |     |     |
|            | 70   | 80  | 90   | 100   | 110 | 120 |     |
| a149-1.pep | ISGDTLRQKAVNLGDALDGVPGIHASQYGGGASAPVIRGQTRRIKVVLNHHGETGDMADF   | :     :     :     :     :     :     :     :     :     :     : | ISGDTLRQKAVNLGDALDGVPGIHASQYGGGASAPVIRGQTRRIKVVLNHHGETGDMADF   | :     :     :     :     :     :     :     :     :     :     : |     |     |     |
| m149-1     |  | 70  | 80   | 90  | 100 | 110 | 120 |
|            |  |   |  |   |     |     |     |
|            | 130  | 140   | 150  | 160   | 170 | 180 |     |
| a149-1.pep | SPDHAIMVDSALSQQVEILRGPVTLLYSSGNVAGLVDVADGKIPFKMPENGVSSELGLRL   | :     :     :     :     :     :     :     :     :     :     : | SPDHAIMVDTALSQQVEILRGPVTLLYSSGNVAGLVDVADGKIPFKMPENGVSSELGLRL   | :     :     :     :     :     :     :     :     :     :     : |     |     |     |
| m149-1     |  | 130   | 140  | 150   | 160 | 170 | 180 |
|            |  |   |  |   |     |     |     |
|            | 190  | 200   | 210  | 220   | 230 | 240 |     |
| a149-1.pep | SSGNLEKLTSGGINIGLGKNFVLHTEGLYRKSGDYAVPRYRNLKRLPDSHADSQTGSIGL   | :     :     :     :     :     :     :     :     :     :     : | SSGNLEKLTSGGINIGLGKNFVLHTEGLYRKSGDYAVPRYRNLKRLPDSHADSQTGSIGL   | :     :     :     :     :     :     :     :     :     :     : |     |     |     |
| m149-1     |  | 190   | 200  | 210   | 220 | 230 | 240 |
|            |  |   |  |   |     |     |     |
|            | 250  | 260   | 270  | 280   | 290 | 300 |     |
| a149-1.pep | SWVGEKGFIGAAYSDRRDQYGLPAHSHEYDDCHADIWQKSLINKRYLQLYPHLLTEEDI    | :     :     :     :     :     :     :     :     :     :     : | SWVGEKGFIGVAYSDRRDQYGLPAHSHEYDDCHADIWQKSLINKRYLQLYPHLLTEEDI    | :     :     :     :     :     :     :     :     :     :     : |     |     |     |
| m149-1     |  | 250   | 260  | 270   | 280 | 290 | 300 |
|            |  |   |  |   |     |     |     |
|            | 310  | 320   | 330  | 340   | 350 | 360 |     |
| a149-1.pep | DYDNPGLSCGFHDDDAHAHAHNKGPKWIQLRNKRYELRAEWKQPFPGFEALRVHLNNDY    | :     :     :     :     :     :     :     :     :     :     : | DYDNPGLSCGFHDDNAHAHTSGRPWIQLRNKRYELRAEWKQPFPGFEALRVHLNNDY      | :     :     :     :     :     :     :     :     :     :     : |     |     |     |
| m149-1     |  | 310   | 320  | 330   | 340 | 350 | 360 |
|            |  |   |  |   |     |     |     |
|            | 370  | 380   | 390  | 400   | 410 | 420 |     |
| a149-1.pep | RHDEKAGDAVENFFNNQTQNARIELRHQPIGRGLKGSGWVQYLGQKSSALSATSEAVKQPM  | :     :     :     :     :     :     :     :     :     :     : | RHDEKAGDAVENFFNNQTQNARIELRHQPIGRGLKGSGWVQYLGQKSSALSATSEAVKQPM  | :     :     :     :     :     :     :     :     :     :     : |     |     |     |
| m149-1     |  | 370   | 380  | 390   | 400 | 410 | 420 |
|            |  |   |  |   |     |     |     |
|            | 430  | 440   | 450  | 460   | 470 | 480 |     |
| a149-1.pep | LLDNKVQHYSFFGVEQANWDNFTLEGGVRVKEKQKASIRYDQKALIDRENYYNHPLPDGLAH | :     :     :     :     :     :     :     :     :     :     : | LLDNKVQHYSFFGVEQANWDNFTLEGGVRVKEKQKASIRYDQKALIDRENYYNHPLPDGLAH | :     :     :     :     :     :     :     :     :     :     : |     |     |     |
| m149-1     |  | 430   | 440  | 450   | 460 | 470 | 480 |
|            |  |   |  |   |     |     |     |
|            | 490  | 500   | 510  | 520   | 530 | 540 |     |
| a149-1.pep | RQTARSFALSGNWYFTPQHKLSLTASHQERLPSTQELYAHGKHVATNTFEVGNKHLNKER   | :     :     :     :     :     :     :     :     :     :     : | RQTARSFALSGNWYFTPQHKLSLTASHQERLPSTQELYAHGKHVATNTFEVGNKHLNKER   | :     :     :     :     :     :     :     :     :     :     : |     |     |     |
| m149-1     |  | 490   | 500  | 510   | 520 | 530 | 540 |
|            |  |   |  |   |     |     |     |
|            | 550  | 560   | 570  | 580   | 590 | 600 |     |
| a149-1.pep | SNNIELALGYEGDRWQYNLALYRNRFGNYIYAQTLNDGRGPKSIEDDSEM KLVRYNQSGA  | :     :     :     :     :     :     :     :     :     :     : | SNNIELALGYEGDRWQYNLALYRNRFGNYIYAQTLNDGRGPKSIEDDSEM KLVRYNQSGA  | :     :     :     :     :     :     :     :     :     :     : |     |     |     |
| m149-1     |  | 550   | 560  | 570   | 580 | 590 | 600 |
|            |  |   |  |   |     |     |     |
|            | 610  | 620   | 630  | 640   | 650 | 660 |     |

|            |  |
|------------|--|
| a149-1.pep | DFYGAEGEIYFKPTPRYRIGVSGDYVRGRLKNLPSLPGREDAYGNRPLIAQADQNAPRVP |
| m149-1     | DFYGAEGEIYFKPTPRYRIGVSGDYVRGRLKNLPSLPGREDAYGNRPFIAQDDQNAPRVP |
|            | 610 620 630 640 650 660                                      |
| a149-1.pep | 670 680 690 700 710 720                                      |
| m149-1     | AARLGVLKASLTDRIDANLDYYRVFAQNKLARYETRTPGHMHMLNLGANYRRNTRYGEWN |
|            | 670 680 690 700 710 720                                      |
| a149-1.pep | 730 740 750 759  |
| m149-1     | WYVKADNLLNQSVAHSSFLSDTPQMGRSFTGGVNVKFX                       |
|            | 730 740 750  |

The following partial DNA sequence was identified in *N.gonorrhoeae* <SEQ ID 607>:

g150.seq (partial)

```

1 .. TACTGCAAGG CAGACCCCTT TCCCGCCGCC CTGCTGCCA ATCAGAAAAT
51 CACCGCCCGC CAATCCGATA AAGACGTGCG CCACATCGAA ATCGATTGAA
101 GCGGTTCGGA TTTGCACTAC CTCCCCGGCG ACGCGCTCGG CGTTTGGTTT
151 GACAACGATC CGGCACTGGT CGGGGAAATC CTAGACCTGC TCGGCATCAA
201 TCCGGCAACG GAAATACAGG CGGGCGGAA AACCCCTGCCG GTTGCCTCCG
251 CACTGTTATC CCATTCGAA CTCACGCAA ACACCCCCGC CTTTGTCAAA
301 GGCTATGCCA CGTTCGCCGA TAATGACGAA CTCGACCGTA TTGCTGCCGA
351 CAACGCCGT TTGCAAGGCT TTGTGCAAAG CACGCCGATT GCCGGTGTGC
401 TGCAACCGTT CCCGGCAAAA CTGACGGCGG AACAAATCGC CGGCCTGCTG
451 CGCCCGCTTG CGCCGCGCCT GTATTGATT TCCTCGTCG AGGCAGGAAGC
501 GGGGGACGAA GTGCACCTGA CCGTCGGCGC AGTGCCTTTC AACACGAAG
551 GGCACGCCAG GGCACGCCAG GCATCGGGTT TCTTGGCGA CGGCGTGGAA
601 GAGGACGGCA CGGTGGCGT GTTTGCGGAA CGCAACGACG GCTTCAGGCT
651 GCCCGAAGAC AGCCGCAAGC CGATTGTGAT GATCGGCTCC GGTACCGGGCG
701 TCGCACCGTT CCGCGCTTC GTCCAACAAC GTGCCGCAGA AAATGCGGAA
751 GGCAGAAAAT GGCTGATTTT CGGCAATCCG CATTGGCCG CGACTTCCT
801 CTATCAGACC GAATGGCAGC AGTTTGCCAA AGACGGCTTC CTGCACAGAT
851 ATGACTTCGC CTGGTCGCGC GATCAGGAAG AAAAATCTA TGTGCAGGAC
901 AAAATCCGCG AACAGGGCGA AGGACTTTGG CAATGGCTGC AGGAAGGCGC
951 GCATATCTAT GTGTGGCGC ATGCGGCAA AATGGCAAA GAAGTGGAAAG
1001 CCGCCTTGCT GGATGTGATT ATCGGGGCAG GGCATTCCGA CGAAGACGGC
1051 GCAGAAGGAT ATTGGATAT GCTGCGCGAA GAAAACGCT ATCAGCGTGA
1101 TTTTATTGA

```

This corresponds to the amino acid sequence <SEQ ID 608; ORF 150.ng>:

g150.pep (partial)

```

1 .. YCKADPFPA LLANQKITAR QSDKDVRHIE IDLSGSDLHY LPGDALGVWF
51 DNDPALVGEI LDLLGINPAT EIQQAGGKTLV VASALLSHFE LTQNTPAFK
101 GYATFADNDE LDRIAADNAV LQGFVQSTPI AGVLHRFPAK LTAEQFAGLL
151 RPLAPRLYSI SSSQAEAGDE VHLTVGAVRF EHEGRARAGG ASGFFADRLE
201 EDGTVRVFAE RNDGFLPED SRKPIVMIGS GTGVAPFRAF VQQRAAENAE
251 GRNWLIIFGNP HFAADFPLYQT EWQQFAKDFL LHRYDFAWSR DQEEDIYVQD
301 KIREQAEGLW QWLQEGAHIY VCGDAAKMAK EVEAALLDVI IGAGHSDEDG
351 AEGYLDMLRE EKRYQRDVY*

```

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 609>:

m150.seq

```

1 ATGCAGAACAA CAAATCCGCC ATTACCGCCT CTGCCGCCCG AAATCACGCA
51 GCTCCCTGTCG GGGCTGGACG CGGCACAATG GGCGTGGCTG TCCGGCTACG
101 CTTGGCAAA AGCAGGAAAC GGGGCATCTG CAGGACTGCC CGCGCTTCAG
151 ACGGCATTGC CGGCAGCAGA ACCTTTTCC GTAACCGTCC TTTCCGCCCTC
201 GCAAACCGGC AATGCGAAAT CCGTTGCCGA CAAAGCGGGC GACAGCCTGG
251 AAGCCGCCGG CATCCAAGTC AGTCGCGCCG AACTGAAAGA CTATAAGGCG
301 AAAAACATCG CCGCGAACG CCGCCTGCTG CTGGTTACCT CCACCCAAGG
351 CGAAGGCAGA CCGCCGAAAG AAGCCGTCGT GCTGCACAAA CTGCTGAACG
401 GCAAAAAAGC CCCGAAATTG GACAAACTCC AATTGCGCGT ACTGGGTTG

```